Figure 1(A)

G <i>GATC</i> GTCTCAGGTCAG <u>CGGAGGGA</u>	25
SL33	
GGAGACTTA TAGACCTATCCAGTCT	50
TCAAGGTGCTCCAGAAAGCAGGAGT	75
TGAAGACCTGGGTGTGAGGGACACA	100
TACATCCTAAAAGCACCACAGCAGA	125
GGAGGCCCAGGCAGTGCCAGGAGTC	150
AAGGTTCCCAGAAGACAAACCCCCT	175
AGGAAGACAGGCGACCTGTGAGGCC	200
CTAGAGCACCACCTTAAGAGAAGAA	225
SL34	
GAGCTGTAAGCCGGCCTTTGTCAGA	250
GCCATC ATG GGGGACAAGGATATGC	275
CTACTGCTGGGATGCCGAGTCTTCT	300
CCAGAGTTCCTCTGAGAGTCCTCAG	325
AGTTGTCCTGAGGGGGAGGACTCCC	350
AGTCTCCTCCAGATTCCCCAGAG	375
TTCTCCTGAGAGCGACGACACCCTG	400
TATCCTCTCCAGAGTCCTCAGAGTC	425
GTTCTGAGGGGGAGGACTCCTCG GA	450
TC CTCTCCAGAGACCTCCTGAGGGG	475
AAGGACTCCCAGTCTCCTCTCCAGA	500
TTCCCCAGAGTTCTCCTGAGGGCGA	525
CGACACCCAGTCTCCTCTCCAGAAT	550
TCTCAGAGTTCTCCTGAGGGGAAGG	575
ACTCCCTGTCTCCTCTAGAGATTTC	600
TCAGAGCCCTCCTGAGGGTGAGGAT	625
GTCCAGTCTCCTCTGCAGAATCCTG	650
CGAGTTCCTTCTTCTCCTCTGCTTT	675
ATTGAGTATTTTCCAGAGTTCCCCT	700

Figure 1(B)

GAGAGAACTCAGAGTACTTTTGAGG	725
GTTTTCCCCAGTCTCCTCTCCAGAT	750
TCCTGTGAGCTCCTCCTCCTCC	775
ACTTTATTGAGTCTTTTCCAGAGTT	800
CCCCTGAGAGAACTCAGAGTACTTT	825
TGAGGGTTTTCCCCAGTCTCTTCTC	850
CAGATTCCTATGACCTCCTCCTTCT	875
CCTCTACTTTATTGAGTATTTTCCA	900
GAGTTCTCCTGAGAGTGCTCAAAGT	925
ACTTTTGAGGGTTTTCCCCAGTCTC	950
CTCTCCAGATTCCTGGGAGCCCCTC	975
CTTCTCCTCCACTTTACTGAGTCTT	1000
TTCCAGAGTTCCCCTGAGAGAACTC	1025
ACAGTACTTTTGAGGGTTTTCCCCA	1050
GTCTCCTCCAGATTCCTATGACC	1075
TCCTCCTTCTCCTCTACTTTATTGA	1100
GTATTTTCCAGAGTTCTCCTGAGAG	1125
TGCTCAAAGTACTTTTGAGGGTTTT	1150
CCCCAGTCTCCTCTCCAGATTCCTG	1175
GGAGCCCCTCCTTCTCCTCCACTTT	1200
ACTGAGTCTTTTCCAGAGTTCCCCT	1225
GAGAGAACTCACAGTACTTTTGAGG	1250
GTTTTCCCCAGTCTCCTCTCCAGAT	1275
TCCTATGACCTCCTCCTCTCT	1300
ACTTTATTGAGTATTTTACAGAGTT	1325
CTCCTGAGAGTGCTCAAAGTGCTTT	1350
TGAGGGTTTTCCCCAGTCTCCTCTC	1375
CAGATTCCTGTGAGCTCCTCTTTCT	1400

Figure 1(C)

CCTACACTTTATTGAGTCTTTTCCA	1425
GAGTTCCCCTGAGAGAACTCAGAGT	1450
ACTTTTGAGGGTTTTCCCCAGTCTC	1475
CTCTCCAGATTCCTGTGAGCTCCTC	1500
CTCCTCCTCCACTTTATTGAGT	1525
CTTTTCCAGAGTTCCCCTGAGTGTA	1550
CTCAAAGTACTTTTGAGGGTTTTCC	1575
CCAGTCTCCTCCAGATTCCTCAG	1600
AGTCCTCCTGAAGGGGAGAATACCC	1625
ATTCTCCTCTCCAGATTGTTCCAAG	1650
TCTTCCTGAGTGGGAGGACTCCCTG	1675
TCTCCTCACTACTTTCCTCAGAGCC	1700
CTCCTCAGGGGGAGGACTCCCTATC	1725
TCCTCACTACTTTCCTCAGAGCCCT	1750
CCTCAGGGGGAGGACTCCCTGTCTC	1775
CTCACTACTTTCCTCAGAGCCCTCA	1800
GGGGAGGACTCCCTGTCTCCTCAC	1825
TACTTTCCTCAGAGCCCTCCTCAGG	1850
GGGAGGACTCCATGTCTCCTCTA	1875
CTTTCCTCAGAGTCCTCTTCAGGGG	1900
GAGGAATTCCAGTCTTCTCTCCAGA	1925
GCCCTGTGAGCATCTGCTCCTC	1950
CACTCCATCCAGTCTTCCCCAGAGT	1975
TTCCCTGAGAGTTCTCAGAGTCCTC	2000
CTGAGGGGCCTGTCCAGTCTCCTCT	2025
CCATAGTCCTCAGAGCCCTCCTGAG	2050
GGGATGCACTCCCAATCTCCTCTCC	2075
AGAGTCCTGAGAGTGCTCCTGAGGG	2100

;;

Figure 1(D)

GGAGGATTCCCTGTCTCCTCCAA	2125
ATTCCTCAGAGTCCTCTTGAGGGAG	2150
AGGACTCCCTGTCTTCTCTCCATTT	2175
TCCTCAGAGTCCTCCTGAGTGGGAG	2200
GACTCCCTCTCTCCACTTTC	2225
CTCAGTTTCCTCCTCAGGGGGAGGA	2250
CTTCCAGTCTTCTCTCCAGAGTCCT	2275
GTGAGTATCTGCTCCTCCTCCACTT	2300
CTTTGAGTCTTCCCCAGAGTTTCCC	2325
TGAGAGTCCTCAGAGTCCTCCTGAG	2350
GGGCCTGCTCAGTCTCCTCTCCAGA	2375
GACCTGTCAGCTCCTTCTTCTCCTA	2400
CACTTTAGCGAGTCTTCTCCAAAGT	2425
TCCCATGAGAGTCCTCAGAGTCCTC	2450
CTGAGGGGCCTGCCCAGTCTCCTCT	2475
CCAGAGTCCTGTGAGCTCCTTCCCC	2500
TCCTCCACTTCATCGAGTCTTTCCC	2525
AGAGTTCTCCTGTGAGCTCCTTCCC	2550
CTCCTCCACTTCATCGAGTCTTTCC	2575
AAGAGTTCCCCTGAGAGTCCTCTCC	2600
AGAGTCCTGT GATC TCCTTCTCCTC	2625
CTCCACTTCATTGAGCCCATTCAGT	2650
GAAGAGTCCAGCAGC <u>CCAGTAGATG</u>	2675
SL26	
<u>AATATACAAGTT</u> CCTCAGACACCTT	2700
GCTAGAGAGTGATTCCTTGACAGAC	2725
AGCGAGTCCTTGATAGAGAGCGAGC	2750
CCTTGTTCACTTATACACTGGATGA	2775
AAAGGTGGACGAGTTGGCGCGGTTT	2800

:

Figure 1(E)

CTTCTCCTCAAATATC <u>AAGTGAAGC</u>	2825
SL27	
<u>AGCCTATCA</u> CAAAGGCAGAGATGCT	2850
GACGAATGTCATCAGCAGGTACACG	2875
GGCTACTTTCCTGT <i>GATC</i> TTCAGGA	2900
AAGCCCGTGAGTTCATAGAGATACT	2925
TTTTGGCATTTCCCTGAGAGAAGTG	2950
GACCCTGATGACTCCTATGTCTTTG	2975
TAAACACATTAGACCTCACCTCTGA	3000
GGGGTGTCTGAGTGATGAGCAGGGC	3025
ATGTCCCAGAACCGCCTCCTGATTC	3050
TTATTCTGAGTATCATCTTCATAAA	3075
GGGCACCTATGCCTCTGAGGAGGTC	3100
ATCTGGGATGTGCTGAGTGGAATAG	3125
GGGTGCGTGCTGGGAGGAGCACTT	3150
TGCCTTTGGGGAGCCCAGGGAGCTC	3175
CTCACTAAAGTTTGGGTGCAGGAAC	3200
ATTACCTAGAGTACCGGGAGGTGCC	3225
CAACTCTTCTCCTCCTCGTTACGAA	3250
TTCCTGTGGGGTCCAAGAGCTCATT	3275
CAGAAGTCATTAAGAGGAAAGTAGT	3300
AGAGTTTTTGGCCATGCTAAAGAAT	3325
ACCGTCCCTATTACCTTTCCATCCT	3350
CTTACAAGGATGCTTTGAAAGATGT	3375
GGAAGAGAGCCCAGGCCATAATT	3400
GACACCACAGATGATTCGACTGCCA	3425
CAGAAAGTGCAAGCTCCAGTGTCAŢ	3450
GTCCCCAGCTTCTCTTC TGA GTGA	3475
AGTCTAGGGCAGATTCTTCCCTCTG	3500

Figure 1(F)

AGTTTGAAGGGGGCAGTCGAGTTTC	3525
TACGTGGTGGAGGGCCTGGTTGAGG	3550
CTGGAGAGAACACAGTGCTATTTGC	3575
ATTTCTGTTCCATATGGGTAGTTAT	3600
GGGGTTTACCTGTTTTACTTTTGGG	3625
TATTTTCAAATGCTTTTCCTATTA	3650
ATAACAGGTTTAAATAGCTTCAGAA	3675
TCCTAGTTTATGCACATGAGTCGCA	3700
CATGTATTGCTGTTTTTCTGGTTTA	3725
AGAGTAACAGTTTGATATTTTGTAA	3750
AAACAAAAACACACCCAAACACACC	3775
ACATTGGGAAAACCTTCTGCCTCAT	3800
TTTGTGATGTGTCACAGGTTAATGT	3825
GGTGTTACTGTAGGAATTTTCTTGA	3850
AACTGTGAAGGAACTCTGCAGTTAA	3875
ATAGTGGAATAAAGTAAAGGATTGT	3900
TAATGTTTGCATTTCCTCAGGTCCT	3925
TTAGTCTGTTGTTCTTGAAAACTAA	3950
AGATACATACCTGGTTTGCTTGGCT	3975
TACGTAAGAAAGTAGAAGAAAGTAA	4000
ACTGTAATAAATAAAAAAAAAAAAA	4025
AAAAA	4031

::

and the state of t

FIG. 2(A)

27	115		70 158	258	188	28 286	386	431
CTCAGGTCAGCGGAGGGAGG	//ctg gagctccaggaaccaggcag lgaggccttggtc tgagacagtatcctcaggtc ////cag GTGCTCCAGAAAGCAGGAGT TGAAGACCTGGGTGTGAGGG ACACATACATACTAAAAGCA intron 1exon II	acaggacacataggactcca gaaacagtgtcagacctggc intron∏exon II	I cacticcicciicaglGTTTT I cgcgicciiciacaglGTTCC intron I <u>II e</u> xon III	gcctttgttagagtctccaa GCCTTTGTCAGAGCCATCAT	intron IT exon III ctcc-cagGCCTGTGGGTC TTCATTG-CCCAGCTCCTGC CCACACTCCTGCCTGCTGCC ctccctcagtcctgtgggat cccatcatacctattcgtgt tcacacgtttacctgctgct	A L G L V C V GCCCTGGCCTGGTGTGT alactggagitggtagatgc	Q A A T S S S S P L V L G T L E E V P T A G S T D P P Q S P Q G A GCAGGCTGCCACCACCACCACCTCCTCCTCTGGGGAGGGGGGGG	lice telitice a cililitat
exon I intron I G C A A Glg t g a g //	lg a g g c c l l g g l · · · · · · c TGAAGACCTGGGTGTGAGGG	caccaagggcccacclgcc laccaagggccgtacccca	tgtaccctga-gtaccctct tgtgccccgaggtgctttct exon ITintron II	AGGAGAAGATCTIglaagtag AGAAGAAGAGCTGTAAGCCG	TTCATTG-CCCAGCTCCTGC cccalcalacclallcglgl	H C K P E E A L E A Q Q E GCACTGCAAGCCTGAGGAAG CCCTTGAGGCCCAACAAGAG acgcccagctttgagcaag gcttccagaaggcaattttc	A G S T D P P TGCTGGGTCAACAGATCCTC	Q R·Q P ACAGAGGC AACCCgalgaggaagaagag gaagclicciccattitctc ticctctitccactlitta
GACCCAGGCTCTGTGAGGAG	gagotocaggaaccaggoag IGTGCTCCAGAAAGCAGGAGT Lexon II	ttgccctgaatgca gagtgcacgacctgactgtg Iron II	legae-eletgetggeegge teatggetetgeetgeeage	CCACAGAGGAGCACCA CCCTAGAGCACCACCTTAAG	intron II exon III ctcc-cagGCCTGTGGGTC ctccclcaglcclglgggat	H C K P E E A GCACTGCAAGCCTGAGGAAG acgcccagclllgagcaag	T L E E V P T ACCCTGGAGGAGGTGCCCAC	P AACCC
CCATTCTGAGGGACGGCGTA GAGTTCGGCCGAAGGAACCT GACCCAGGCTCTGTGAGGAG	TCTTCAAG g11///cag exon 1 introl	acagagcagaggalgcacag gglglgccagcaglgaalgt ttgccclgaalgca caccaagggccccacclgco CCACAGCAGAGGAGGCCCAG GCAGTGCCAGGAGTCAAGgt gagtgcacgacctgactgtg laccaagggccgtacccca exon III intron II	cagagicigocotcacotco etacigicagicotglagaa icgac-olotgotggcoggo igiacoolga-glacoolot agcacoggcoolglagcoac coacigicalicotggigoo icatggolotgoolgocago igigocoogaggigoiitel exon IlTintron II	CAGGGGACAGGCCAACCCAG AGGACAGGATTCCCTGGAGG CAGAAGACAAACCCCCTAGG AAGACAGGCGACCTGTGAGG	gcololoacacaclocolot gcoaclggcaclglocolot	S L·E Q R S L TCTCTTGAGCAGAGGAGTCT cclcltllclaaacctlcc	S P L V L G CCTCTCCTCTGGGC	S A F P T T I N F T R Q R Q P TCCGCCTTTCCCACTACCAT CAACTTCACTCGACAGAGGC AACCC
CCATTCTGAGGGACGGCGTA	AGACTTATAGACCTATCCAG TCTTCAAGg1/	acagagcagaggalgcacag CCACAGCAGAGGAGGCCCAG	cagagletggeeteaeetee ageaeeggeeetglageeae	CAGGGGACAGGCCAACCCAG CAGAAGACAAACCCCCTAGG	l gglicag-tictcagctgag i GGgtgagtitctcagctgag on III G_intron III	M S L · E Q R S L · CCT:GACGAGAGTCATCATG TCTTTGAGCAGAGGAGTCT cctgaacaatattcatcatg cctcttttctaaaccttcc	O A A T S S S S P L GCAGGCTGCCACCTCCTCT CCTCTCCTC agaggatcccca	
					0	***		77.72

FIG. 2(B)

294 ಽಽ lccctcglcclcclcclgf tittctitclcatcclcatc cicctcicigcitcigcgtt ciccagGGGACAAGGATATG CCTACTGCTGGGATGCCGAG intron IIII D K D M P T A G M P S intron III D K D M exon IV

394 46

494 ACCCTGTATCCTCTCCAGAG TCCTCAGAGTCGTTCTGAGG GGGAGGACTCCTCGGATCCT CTCCAGAGACCTCCTGAGGG GAAGGACTCCCAGTCTCCTC

594 TCCAGATTCCCCAGAGITCT CCTGAGGGCGACGCCCA GTCTCCTCCAGAATTCTC AGAGTTCTCCTGAGGGGAAG GACTCCCTGTCTCCTCAGA

Q | P | Q S S P E G D D T Q S P L Q N S O S S P E G K D S L S P L E

694 GATITOTCAGAGCCCTCCTG AGGGTGAGGATGTCCAGTCT CCTCTGCAGAATCCTGCGAGG TTCCTTCTTCTCCTGCTT TATTGAGTATTTTCCAGAGT

794 180 TCCCTGAGAGTATTCAAAG TCCTTTTGAGGGTTTTCCCC AGTCTGTTCTCCAGATTCCT GTGAGCGCCGCCTCCTCCTC CACTTTAGTGAGTATTTTCC

AGAGTICCCCTGAGAGTACT CAAAGTCCTTTTGAGGGTTT TCCCCAGTCTCCAGA TTCCTGTGAGCGCTCCTTC TCCTCCACTTTATTGAGTAT S S T L L S 1. P V S R S F TITCCAGAGITCCCCTGAGA GAAGTCAGAGAACTICTGAG GGTITTGCACAGTCTCCTCT CCAGATTCCTGTGAGCTCCT CCTCGTCCTCCACTTTACTG

	A1				
	CI AGICTITICCAGAGITCCCC TGAGAGAACTCAGAGIACTI TTGAGGGITITCCCCAGGICT CCACTCCAGATICCTGIGAG CCGCTCCTTCTCCTCCACTT 1094	GTACTT TTGAGGGTTTTCCCCAGTCT CC	ACTCCAGATTCCTGTGAG	CCGCTCCTCCCCCACTT 10	1094
- -	C1 TATTGAGTATTTTCCAGAGT TCCCCTGAGAGAACTCAGAG TACTTTTGAGGTTTTGCCC AGTCTCCTCCCAGATTCCT GTGAGCTCCTCCTCCTCT 1194 L S L F O S P E R T O S T F E G F A Q S P L O I P V S S S S S 313	CAGAG TACTTTGAGGGTTTTGCCC AG	TCTCCTCCCAGATTCCT	GTGAGCTCCTCCTC 11	1194 313
7 0	CI CACTITATIGAGICTITICC AGAGITCCCCIGAGAGAACT CAGAGIACTITIGAGGGITI TCCCCAGICTCTCCCAGA TICCTATGACCTCCTTC 1294	GAACT CAGAGTACTTTTGAGGGTTT TCC	CCCAGTCTCTCCCAGA	TICCTATGACCTCCTC 128	294 346
∢ O	C1 TCCTCTACTITATIGAGTAT TITCCAGAGTTCTCCTGAGA GTGCTCAAAGTACTTTTGAG GGTTTTCCCCAGTCTCCTCT CCAGATTCCTGGGAGCCCCT 1394 S S T L L S L F O S P E S A O S P L O 1 P G S P S 380	IGAGA GIGCTCAAAGTACTTTTGAG GGT	TITICCCAGICICCICT	CCAGATTCCTGGGAGCCCCT 136	330
< ∪	CI CCTICTCCTCCACATITACTG AGTCTTTTCCAGAGTTCCCC TGAGAGCTCACAGTACTT TTGAGGGTTTTCCCCAGTCT CCTCTCCAGATTCCTATGAC 1494	CCCC TGAGAGCTCACAGTACTT TTG	AGGGTTTTCCCCAGTCT (CCTCTCCAGATTCCTATGAC 149	194 13
ΑÖ	CI CTCCTCCTTCTCCTACTT TATTGAGTATTTTACAGAGT TCTCCTGAGAGTGCTCAAAG TGCTTTTGAGGGTTTTCCCC AGTCTCCTCTCC	GAGT TCTCCTGAGAGTGCTCAAAG TGC	TTTTGAGGGTTTTCCCC / F E G F P Q	AGTCTCTCCAGATTCCT 159	94 46
₹5	CI GTGAGCTCCTCTTTCTCCTA CACTITATIGAGTCTTTTCC AGAGTTCCCCTGAGAGACT CAGAGTACTTTTGAGGGTTT TCCCCAGTCTCTCCTCCAGA 1694	TICC AGAGITCCCTGAGAGACT CAG	AGTACTTTTGAGGGTTT T	CCCCAGTCTCTCTCCAGA 1694	30 %
A D	CI ITCCTGTGAGCTCCTCCTCC TCCTCCTCCACTITATTGAG TCTTTTCCAGAGTTCCCCTG AGTGTACTCCAAAGTACTTTT GAGGGTTTTCCCCAGTCTCC 1794	IGAG TCTTTTCCAGAGTTCCCCTG AGTC	STACTCAAAGTACTTTT G	AGGGTTTTCCCCAGICTCC 1794 G F P Q S P 513	4 E

•

1894	1994	2094 613	2194 646	2294 680	2394	2494	2594 780	2694 813 :;
TCTCCAGATTCCTCAGAGTC CTCCTGAAGGGGAGAATACC CATTCTCCTCCAGATTGT TCCAAGTCTTCCTGAGTGGG AGGACTCCCTGTCTCCTCAC L Q P Q S P E G E N T H S P L Q I V P S L P E W E D S L S P H	CTCCCTGTCTCCTCACTACT S L S P. H Y F						TCCTCTCCAGAGACCTGTCA 2	SAGTCTCTCCCAGAGTCC 2
T TCCAAGTCTTCCTGAGTGGG	TACTITCCTCAGAGCCCTCC TCAGGGGGGGGGGGGGGGGGGGGG	TTCCTCAGAGCCCTCAGGGG GAGGACTCCCTGACTTTCCTCAGAGCCCTC CTCAGGGGGGGGGG	GAGTCCTCTTCAGGGGGGGG AATTCCAGTCTTCTCTCCAG AGCCCTGTGAGCATCTGCTC CTCCTCCACTCCA	AGTICICAGAGICCICCTGA GGGGCCTGTCCAGTCTCCTCAGAGCCCT CCTGAGGGGATGCACTCCCA ATCTCCTCTCC	AGAGTGCTCCTGAGGGGGGAG GATTCCTCTCCA AATTCCTCAGAGTCCTCTTG AGGGAGAGGGACTCCCTGTCT TCTCTCCATTTTCCTCAGAG	TCCTCCTGAGTGGGAGGACT CCCTCTCTCTCCTCTCTCTCAGGG GGAGGACTTCCAGTTCTC TCCAGAGTCTGTGAGTATC P P E W E D S L S P L H F P Q F P Q G E D F Q S S L Q S P V S I	TGCTCCTCCTCCTCCTTT GAGTCTTCCCCAGAGTTTCC CTGAGAGTCCTCAGAGTCCT CCTGAGGGGCCTGCTCAGTC TCCTCTCCAGAGACCTGTCA	TTCCCATGAGAGTCCTCAGAGGGGCCTGCC CAGTCTCCTCTCC
C CATTCTCCTCCAGATTG	T CTCCTCACTACTTTCCTCAG	CTACTTTCCTCAGAGCCCTC	AGCCCTGTGAGCATCTGCTC	TCCATAGTCCTCAGAGCCCT H S P Q S P	AATTCCTCAGAGTCCTCTTG	CCTCAGTTTCCTCCTCAGGG	CTGAGAGTCCTCAGAGTCCT E S P Q S P	TCCCATGAGAGTCCTCAGA S H E S P 0 S
C CTCCTGAAGGGGAGATAC P P E G E N T	CAGGGGGAGGACTCCCTATO	GAGGACTCCCTGTCTCCTCA E D S L S P H	AATTCCAGTCTTCTCCCAG	GGGGCCTGTCCAGTCTCCTC G P V Q S P L	GATTCCCTGTCTCCTCCA D S L S P L Q	CCCTCTCCTCCCACTTT L S P L H F	GAGTCTTCCCCAGAGTTTCC	GCTCCTTCTTCTCCTACACT TTAGCGAGTCTTCTCCCAAAG
TCTCCAGATTCCTCAGAGT	TACTTTCCTCAGAGCCCTCC	TTCCTCAGAGCCCTCAGGGG	GAGTCCTCTCAGGGGGAGG		AGAGTGCTCCTGAGGGGGAG		TGCTCCTCCACTTCTTT	GCTCCTTCTTCTCCTACACT \$ F F S Y T
Z 2	<u> 5</u> 2	G A	CA	- C A	CA	5 2	A 20	₹5

G

846 C1 IGTGAGCTCCTTCCCCTCCT CCACTTCATCGAGTCTTTCC CAGAGTTCTCCTGTGAGCTC CTTCCCCTCCTCCACTTCAT CGAGTCTTTCCAAGAGTTCC 2794 S F P S S T S S S L S P V S S S 0 S __1 T S S S S | S | P | S |

2894 880 CCCAGTAGATGAATATACAA S E G S S R E E E G P S ... AGTGAGGGTTCCAGCAG CCGTGAAGAGGAGGGGCCAA P V D CI CCTGAGAGICCTCTCCAGAG TCCTGTGATCTCCTTCTCCT CCTCCACTTCATTGAGCCCCA TTCAGTGAAGAGTCCAGCAG S ш S ш SS S. P | V | S F S S S T S L S P O A1

2994 108 523 ACACTGGATGAAAAGGTGGA T L D E K V D GTAATCACTAAGAAGGTGGC GTTCCTCAGACACCTTGCTA GAGAGTGATTCCTTGACAGA CAGCGAGTCCTTGATAGAGA GCGAGCCCTTGTTCACTTAT A A E S L F GCACCTCTTGTATCCTG---A

913

FTY

ــ ــ

ш

ഗ

ш

L S

ш

တ

SLTD

E S D

140 623 3094 946 AATGCTGGAGAGTGTCATCA AAAATTACAAGCACTGTTTT GATGCTGACGAATGTCATCA GCAGGTACACGGGCTACTTT >-N V I S ш M L T _ ≥ CAGCCTATCACAAAGGCAGA Q P I T K A E E P V T K A E GAGCCAGTCACAAAGGCAGA PVTKA IGATITGGTTGGTTTTCTGC TCCTCAAATATCGAGCCAGG CGAGTTGGCGCGGTTTCTTC TCCTCAAATATCAAGTGAAG LKYRAR L K Y Q V K DIVGFLL A 2

174 723 979 CTCCTATGTCTTTGTAAACA 3191 CTCCTATGTCCTTGTCACCT 7 1 N 1 C SYVFVN S CCGTGAGTTCATAGAGATAC TITITGGCATTTCCCTGAGA GAAGTGGACCCT...GATGA
REFIEILFGISLR EVDP.DD GAAGCAGCCCACCGGCCA エ တ **—** с. ۵ E A TCTTTGGCATTGACGTGAAG FGIDVK CTCTGAGTCCTTGCAGCTGG E S L Q L V တ CCTGAGATCTTCGGCAAAGC CCTGTGATCTTCAGGAAAGC PEIFGKA œ Z 2

3291 1012 **ATGATTGCAATGGAGGGCGG ATCATCTTCATAAAGGGCAC** - F - X M I A M GATCATGCCCAAGACAGGCT ICCTGATAATTGTCCTGGTC GGGCATGTCCCAGAACCGCC TCCTGATTCTTATTCTGAGT S 7 | 7 | 1 G F ONBL N D X S W S GGCCTGCTGGGTGATAATCA GGGTGTCTGAGTGAGCA G C L S D E O O N O 9 7 7 9 GCCTAGGTCTCTCCTATGAT CATTAGACCTCACCTCTGAG S Y D ₹ 5

3391 GGAGCACAGTGCCTATGGGG AGCCCAGGAAGCTGCTCACC GGAGCACTITGCCTTTGGGG AGCCCAGGGAGCTCCTCACT P = K L L ш E H S A Y G **ATGGAGGTGTATGATGGGAG** TCTGGGATGTGCTGAGTGGA ATAGGGGTGCGTGCTGGGAG ≃ છ .:1 .≻ ... __ __ TCTGGGAGGAGCTGAGTGTG 7 ^ ₹ CTATGCCTCTGAGGAGGTCA CCATGCTCCTGAGGAGGAAA ш ₹5

1023 3491 1079 TCCAAGGGCCCTCGCTGAAA TCCAAGAGCTCATTCAGAAG œ ۵. CGCTATGAGTTCCTGTGGGG CGTTACGAATTCCTGTGGGG ග ≥ ≥ <u>~</u> >-GTACCTGGAGTACCGGCAGG TGCCGGACAGTGATCCCGCA TTACCTAGAGTACCGGGAGG TGCCCAACTCTTCTCCTCCT ۵. ۵. S ഗ 2 ۳ ۲ > 0 E Y B YLEYR CAAGATTTGGTGCAGGAAAA AAAGTTTGGGTGCAGGAACA ш ж ш G >

307 1123 3591 1112 A L R E E E E GCTTTGAGAGGAGGAAGA GCTTTGAAAGATGTGGAAGA ш Œ TGCAAGAGTTCGCTTTTTCT TCCCATCCCTGCGTGAAGCA TTCCATCCTCTTACAAGGAT ш P S S Y K S L R GAATACCGTCCCTATTACCT ARVRFFF P - T GAGTATGTGATCAAGGTCAG GAGTTTTGGCCATGCTAAA E F L A M L K EYVIKVS CCAGCTATGTGAAAGTCCTT **CCATTAAGAGGAAAGTAGTA**

1223 1142 3691 TCTCTTCTGAGTGAAGTCTA TCCAGGGCCGCGTCCAGCAG CTTCCCCTGCCTCGTGTGAC ဟ CAGTGTCATGTCCCCCAGCT တ ۵. တ Σ > တ GCCACAGAAAGTGCAAGCTC A T E S A S S GGGACTGGGCCAGTGCACCT 36GAGTCTGAGCATGAGTTG CAGCCAAGGCCAGTGGGAGG GAGAGCCCAGGCCATAATTG ACACCACAGATGATTCGACT ဟ 0 0 **⊢** V OPA

TGGTTGAGGCTGGAGAAC ACAGTGCTATTTGCATTTCT GTTTGAAGGGGCAGTCGAG TTTCTACGTGGTGGAGGGCC GTTCTCAGTAGTAG-...ATGAGGCCCATTCTTCA CTCTGAAGAGGGGGTCAGT GGGCAGATTCTTCCCTCTGA

AATGAACTTCAGCATCCAAG AAATAGCTTCAGAATCCTAG TT--TTTTAAGGGATGGTTG GGGTTTACCTGTTTTACTTT TGGGTATTTTTCAAATGCTT TTCCTATTAATAACAGGTTT AGATTIATCTTTGTTCTCTT TTGGAATTGTTCAAATGTTT reftctattgggtgacttgg GTTCCATATGGGTAGTTATG

1476 3991 TCCATTCTATTTTGTGAATT CACCCAAACACACACATTG GTTTTATTCAGATTGGGAAA ATTTTGTAAAAACAAAAACA ITTATGAATGACAGCAGT-C ACACAGTTCTGTGTATATAG TITAAGGGTAAGAGTCTTGT TITATGCACATGAGTAACAGTTTGAT TITATGCACATGAGAGTAACAGTTTGAT

ATTAAGAGATAGTCAATTCT TIGIGAIGICACAGGITA AIGIGGIGITACIGIAGGAA ITTICITGAAACTGIGAAGG AACTCIGCAGTTAAAIAGIG GGGATAATAACAGCAGTGGA ATAAGTACTTAGAATGTGA AAAATGAGCAGTAAAATAGA TGAGATAAAGAACTAAAGAA SGAAAACCTTCTGCCTCATT

AGAGAATTAAATCTGAATA 1676 AATGITIGCATITCCICAGG ICCITIAGICIGIIGIICII GAAAACIAAAGAIACAIACC IGGITIGCIIGGCIIACGIA 4191 IGCCTIATACCTCAGICIAI ICIGIAAAATITITAAAGAI ATAIGCATACCIGGAITICC TIGGCTICITIGAGAAIGTA GAATAAAGTAAAGGATTGTT

E.G.

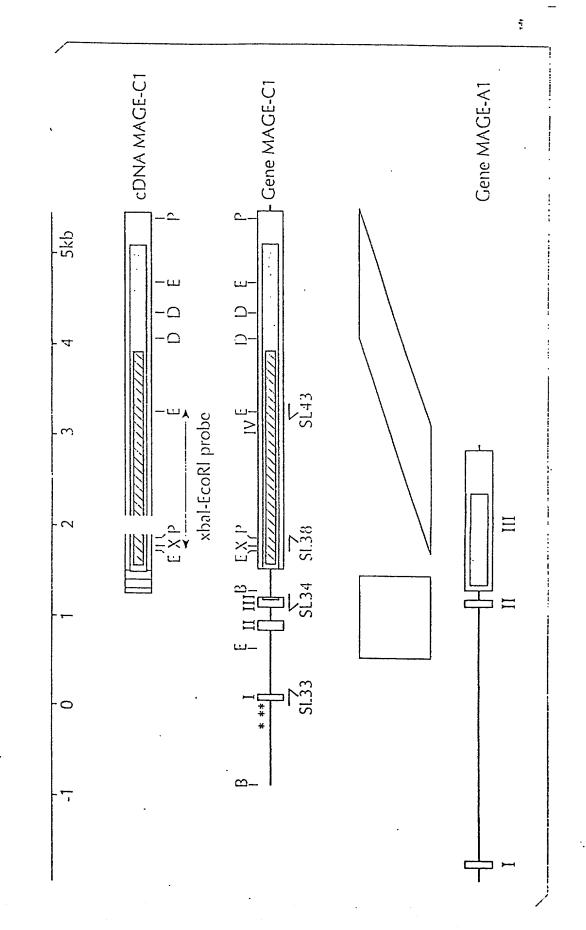
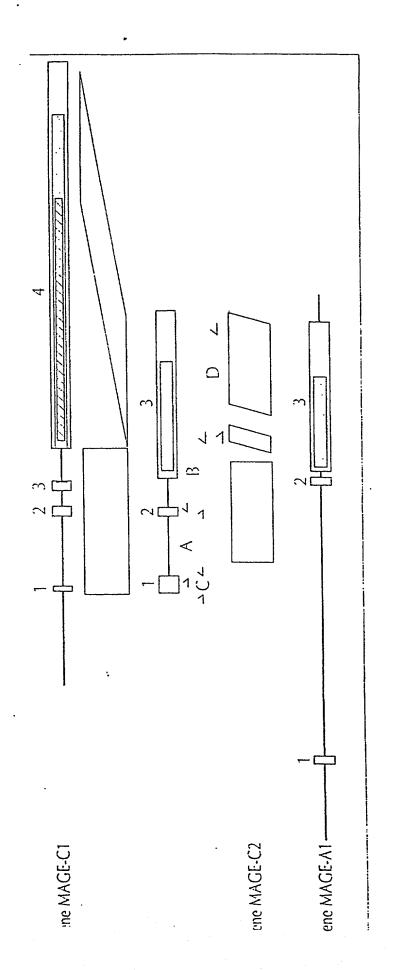


FIG. 4



::

The sequence corresponds to nucleotides 3761 to 4801 of PAC clone 232G24 (GenBank accession number AL022152). It has been renumbered in this figure, taking as nucleotide number 1 the first nucleotide of the start codon.

Primers SL164 and SL165 that were used in the RT-PCR assay are indicated.

${\tt ATGCCTCTTTTCCAAACCTTCCACGCCTCAGCTTTGAGGAAGACTTCCA}$	50
GAACCCGAGTGTGACAGAGGACTTGGTAGATGCACAGGATTCCATAGATG	100
${\bf A} {\tt GGAGGAGGATGCCTCCTCCACTTCCTCTTTCCACTTTTTA}$	150
TTCCCCTCCTCCTTCCTTGTCC <u>TCATCCTCACCCTTGTCCTCA</u> CCCTT sl164	200
ACCCTCTACTCTCATTCTGGGTGTTCCAGAAGATGAGGATATGCCTGCTG	250
CTGGGATGCCACCTCTTCCCCAGAGTCCTCCTGAGATTCCTCCCCAGGGT	300
CCTCCCAAGATCTCTCCCCAGGGTCCTCCGCAGAGTCCTCCCCAGAGTCC	350
TCTAGACTCCTGCTCATCCCCTCTTTTGTGGACCCGATTGGATGAGGAGT	400
CCAGCAGTGAAGAGGAGGATACAGCTACTTGGCATGCCTTGCCAGAAAGT	450
GAATCCTTGCCCAGGTATGCCCTGGATGAAAAGGTGGCTGAGTTGGTGCA	500
GTTTCTTCTCCTCAAATATCAAACAAAAGAGCCTGTCACAAAGGCAGAGA	550
TGCTGACGACTGTCATCAAGAAGTATAAGGACTATTTTCCCATGATCTTC	600
GGGAAAGCCCATGAGTTCATAGAGCTAATTTTTGGCATTGCCCTGACTGA	650
TATGGACCCCGACAACCACTCCTATTTCTTTGAAGACACATTAGACCTCA	700
CCTATGAGGG <u>AAGCCTGATTGATGACCAGGG</u> CATGCCCAAGAACTGTCTC 8L165	750
CTGATTCTTATTCTCAGTATGATCTTCATAAAGGGCAGCTGTGTCCCCGA	800
GGAGGTCATCTGGGAAGTGTTGAGTGCAATAGGGGTGTGTGCTGGGAGGG	850
AGCACTTTATATATGGGGATCCCAGAAAGCTGCTCACTATACATTGGGTG	900
CAGAGAAAGTACCTGGAGTACCGGGAGGTGCCCAACAGTGCTCCTCCACG	950
TTATGAATTTTTGTGGGGTCCAAGAGCCCATTCAGAGGCCAGCAAGAGAA	1000
GTCTTAGAGTTTTTATCCAAGCTATCCAGTATCATCCCTAG	1041

5

25

* \$.	Figure 6 Amino-acid sequence of the putative MAGE-C3 protein (SEQ ID NO: 22)	;
	MPLFPNLPRLSFEEDFQNPSVTEDLVDAQDSIDEEEEDASSTSSSSFHFL	50
5	FPSSSSLSSSSPLSSPLPSTLILGVPEDEDMPAAGMPPLPQSPPEIPPQG	100
	PPKISPQGPPQSPPQSPLDSCSSPLLWTRLDEESSSEEEDTATWHALPES	150
	ESLPRYALDEKVAELVQFLLLKYQTKEPVTKAEMLTTVIKKYKDYFPMIF	200
	GKAHEFIELIFGIALTDMDPDNHSYFFEDTLDLTYEGSLIDDQGMPKNCL	250
	LILILSMIFIKGSCVPEEVIWEVLSAIGVCAGREHFIYGDPRKLLTIHWV	300
10	QRKYLEYREVPNSAPPRYEFLWGPRAHSEASKRSLRVFIQAIQYHP	346

Figure 7 Nucleotide sequence of gene MAGE-B5 (SEQ ID NO: 23)

This sequence is the antiparallel sequence corresponding to nucleotides 123358 to 124185 of BAC clone 466O4 (GenBank accession number AC005297). The sequence has been renumbered in this figure, taking the first nucleotide of the start codon as nucleotide number 1.

Primers SL189 and SL190 that were used in the RT-PCR assay are indicated.

ATGACTTCTGCAGGTGTTTTTAATGCAGGATCTGACGAAAGGGCTAACAG	50
TAGAGATGAGGAGTACCCATGTTCCTCAGAGGTCTCACCCTCCACTGAGA	100
GTTCATGCAGCAATTTCATAAATATTAAGGTGGGTTTGTTGGAGCAGTTC	150
CTGCTCTACAAGTTCAAAATGAAACAGCGTATTTTGAAGGAAG	200
GAAGATTGTCAACCCAAGATACCAAAACCAGTTTGCTGAGATTCACAGAA	250
GAGCTTCTGAGCACATTGAGGTTGTCTTTTGCAGTTGACTTGAAGGAAG	300
AACCCAACTTGTCACTTATATGACCTTGTCAGCAAGCTGAAACTCCCCAA	350
CAATGGGAGGATTCATGTT <u>GGCAAAGTGTTACCCAAGACTGGTC</u> TCCTCA sl189	400
TGACTTTCCTGGTTGTGATCTTCCTGAAAGGCAACTGTGCCAACAAGGAA	450
GATACCTGGAAATTTCTGGATATGATGCAAATATATGATGGGAAGAAGTA	500
CTACATCTATGGAGAGCCCAGGAAGCTCATCACTCAGGATTTCGTGAGGC	550
TAACGTACCTGGAGTACCACCAGGTGCCCTGCAGTTATCCTGCACACTAT	600
CAATTCCTTTGGGGTCCAAGAGCCTATACTGAAACCAGCAAGATGAAAGT	650
CCTGGAATATTTGGCCAAGGTCAATGATATT <u>GCTCCAGGTGCCTTCTCAT</u> SL190	700
<u>CACAA</u> TATGAAGAGGCTTTGCAAGATGAGGAAGAGGCCCAAGCCAGAGA	750
TGCAGCCGAAACTGGCACTACTGCAGTGGCCAAGACTGTCTCAGGGCGAA	800
GTTCAGCAGCTTCTCAACCCTATTGA	828

25

5

10

. *	Figure 8 Amino-acid sequence of the putative MAGE-B5 protein (SEQ ID NO: 24)	•
	MTSAGVFNAGSDERANSRDEEYPCSSEVSPSTESSCSNFINIKVGLLEQF	50
5	LLYKFKMKQRILKEDMLKIVNPRYQNQFAEIHRRASEHIEVVFAVDLKEV	100
	NPTCHLYDLVSKLKLPNNGRIHVGKVLPKTGLLMTFLVVIFLKGNCANKE	150
	DTWKFLDMMQIYDGKKYYIYGEPRKLITQDFVRLTYLEYHQVPCSYPAHY	200
	QFLWGPRAYTETSKMKVLEYLAKVNDIAPGAFSSQYEEALQDEEESPSQR	250
	CSRNWHYCSGQDCLRAKFSSFSQPY	275
10		

This sequence is the antiparallel sequence corresponding to nucleotides 146417 to 147640 of BAC clone 46604 (GenBank accession number AC005297). The sequence has been renumbered in this figure, taking the first nucleotide of the start codon as nucleotide number 1.

Primers SL191 and SL192 that were used in the RT-PCR assay are indicated.

10	ATGCCTCGGGGTCACAAGAGTAAGCTCCGTACCTGTGAGAAACGCCAAGA	50
	GACCAATGGTCAGCCACAGGGTCTCACGGGTCCCCAGGCCACTGCAGAGA	100
Sandry Sandry Berling	AGCAGGAAGAGTC <u>CCACTCTTCCTCATCCTCTTCTCG</u> CGCTTGTCTGGGT SL191	150
	GATTGTCGTAGGTCTTCTGATGCCTCCATTCCTCAGGAGTCTCAGGGAGT	200
150	GTCACCCACTGGGTCTCCTGATGCAGTTGTTTCATATTCAAAATCCGATG	250
TOTAL	TGGCTGCCAACGGCCAAGATGAGAAAAGTCCAAGCACCTCCCGTGATGCC	300
	TCCGTTCCTCAGGAGTCTCAGGGAGCTTCACCCACTGGCTCTCCTGATGC	350
	AGGTGTTTCAGGCTCAAAATATGATGTGGCTGCCAACGGCCAAGATGAGA	400
	AAAGTCCAAGCACTTCCCATGATGTCTCCGTTCCTCAGGAGTCTCAGGGA	450
20	GCTTCACCCACTGGCTCGCCTGATGCAGGTGTTTCAGGCTCAAAATATGA	500
	TGTGGCTGC <u>CGAGGGTGAAGATGAGGAAAGTG</u> TAAGCGCCTCACAGAAAG SL192	550
	CCATCATTTTTAAGCGCTTAAGCAAAGATGCTGTAAAGAAGAAGAAGGCGTGC	600
\$ 100°	ACGTTGGCGCAATTCCTGCAGAAGAAGTTTGAGAAGAAGAGTCCATTTT	650
25	GAAGGCAGACATGCTGAAGTGTGTCCGCAGAGAGTACAAGCCCTACTTCC	700
	CTCAGATCCTCAACAGAACCTCCCAACATTTGGTGGTGGCCTTTGGCGTT	750
	GAATTGAAAGAAATGGATTCCAGCGGCGAGTCCTACACCCTTGTCAGCAA	800
	GCTAGGCCTCCCCAGTGAAGGAATTCTGAGTGGTGATAATGCGCTGCCGA	850
	AGTCGGGTCTCCTGATGTCGCTCCTGGTTGTGATCTTCATGAACGGCAAC	900
30	TGTGCCACTGAAGAGGAGGTCTGGGAGTTCCTGGGGTCTGTTGGGGATATA	950
	TGATGGGATCCTGCATTCAATCTATGGGGATGCTCGGAAGATCATTACTG	1000
	AAGATTTGGTGCAAGATAAGTACGTGGTTTACCGGCAGGTGTGCAACAGT	1050
	GATCCTCCATGCTATGAGTTCCTGTGGGGTCCACGAGCCTATGCTGAAAC	1100
	CACCAAGATGAGAGTCCTGCGTGTTTTTGGCCGACAGCAGTAACACCAGTC	1150
35	CCGGTTTATACCCACATCTGTATGAAGACGCTTTGATAGATGAGGTAGAG	1200
	AGAGCATTGAGACTGAGAGCTTAA	1224

,	Figure 10 Amino-acid sequence of the putative MAGE-B6 protein (SEQ ID NO: 26)	•
	MPRGHKSKLRTCEKRQETNGQPQGLTGPQATAEKQEESHSSSSSSRACLG	50
5	DCRRSSDASIPQESQGVSPTGSPDAVVSYSKSDVAANGQDEKSPSTSRDA	100
	SVPQESQGASPTGSPDAGVSGSKYDVAANGQDEKSPSTSHDVSVPQESQG	150
	ASPTGSPDAGVSGSKYDVAAEGEDEESVSASQKAIIFKRLSKDAVKKKAC	200
	$\verb TLAQFLQKKFEKKESILKADMLKCVRREYKPYFPQILNRTSQHLVVAFGV $	250
	ELKEMDSSGESYTLVSKLGLPSEGILSGDNALPKSGLLMSLLVVIFMNGN	300
10	CATEEEVWEFLGLLGIYDGILHSIYGDARKIITEDLVQDKYVVYRQVCNS	350
	DPPCYEFLWGPRAYAETTKMRVLRVLADSSNTSPGLYPHLYEDALIDEVE	400
	RALRLRA	407